

改图为文远楼测试的日照，将日照错误值改成0

如果改成拉格朗日

聚类算法

针对聚类算法，我们想另外一个思路，把一天当做一个样本，一个样本24个特征值，特征值就是该小时的能耗值。

* 1. 函数数据的清洗

A 我们通过将天气数据变成（0，-1，-2）来表示晴阴雨雪

B 经过观察，发现辐照强度测量值在每天较晚时刻会出现错误值，我们通过将有极多错误值的辐照强度改为同月份内各天同一时刻的辐射测量有效值的平均值来代替。（averageifs）

1.2函数数据的预处理

先归一化，然后方差化消除影响。

Datamat2 不考虑小时，不做归一化centilist3,mynewassments3

Datamat3 考虑小时，不做归一化centilist4,mynewassments4

Datamat4 考虑小时，归一化centilist5,mynewassments5

1.2.1 datamat3计算得到

分为5类。

0.0 : 540

1.0 : 1905

2.0 : 105

3.0 : 270

4.0 : 108

距离质心之和

{0.0: [0.0, 2604.8154145393264], 1.0: [1.0, 8598.47065414216], 2.0: [2.0, 540.2039834945024], 3.0: [3.0, 1122.9539396447826], 4.0: [4.0, 469.2960561444093]}

质心：